



OICE

RAW SEQUENCE LISTING

DATE: 01/16/2002

PATENT APPLICATION: US/10/029,137

TIME: 11:09:32

Input Set : A:\407T-301400US.txt

Output Set: N:\CRF3\01162002\J029137.raw

ENTERED

3 <110> APPLICANT: NGUYEN, MAI
 5 <120> TITLE OF INVENTION: IDENTIFICATION OF A NOVEL ENDOTHELIAL-DERIVED GENE EG-1
 7 <130> FILE REFERENCE: 407T-301400US
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/029,137
 10 <141> CURRENT FILING DATE: 2001-12-12
 12 <160> NUMBER OF SEQ ID NOS: 15
 14 <170> SOFTWARE: PatentIn version 3.0
 16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 1281
 18 <212> TYPE: DNA
 19 <213> ORGANISM: Homo sapiens
 21 <220> FEATURE:
 22 <221> NAME/KEY: CDS
 23 <222> LOCATION: (6)..(539)
 25 <400> SEQUENCE: 1
 26 cagaa atg gcg gct cca cta ggg ggt atg ttt tct ggg cag cca ccc ggt 50
 27 Met Ala Ala Pro Leu Gly Gly Met Phe Ser Gly Gln Pro Pro Gly
 28 1 5 10 15
 30 ccc cct cag gcc ccg ccg gcc ctt ccg gcc caa gct cgg att att cag 98
 31 Pro Pro Gln Ala Pro Pro Gly Leu Pro Gly Gln Ala Ser Leu Leu Gln
 32 20 25 30
 34 gca gct cca gcc gct cct aga cct tcc agc agt act ttg gtg gac gag 146
 35 Ala Ala Pro Gly Ala Pro Arg Pro Ser Ser Ser Thr Leu Val Asp Glu
 36 35 40 45
 38 ttg gag tca tct ttc gag gct tgc ttt gca tct ctg gtg agt cag gac 194
 39 Leu Glu Ser Ser Phe Glu Ala Cys Phe Ala Ser Leu Val Ser Gln Asp
 40 50 55 60
 42 tat gtc aat gcc acc gat cag gaa gaa att cga acc ggt gtt gat cag 242
 43 Tyr Val Asn Gly Thr Asp Gln Glu Glu Ile Arg Thr Gly Val Asp Gln
 44 65 70 75
 46 tgt atc cag aag ttt ctg gat att gca aga cag aca gaa tgt ttt ttc 290
 47 Cys Ile Gln Lys Phe Leu Asp Ile Ala Arg Gln Thr Glu Cys Phe Phe
 48 80 85 90 95
 50 tta caa aaa aga ttg cag tta tct gtc cag aaa cca gag caa gtt atc 338
 51 Leu Gln Lys Arg Leu Gln Leu Ser Val Gln Lys Pro Glu Gln Val Ile
 52 100 105 110
 54 aaa gag gat gtg tca gaa cta agg aat gaa tta cag ccg aaa gat gca 386
 55 Lys Glu Asp Val Ser Glu Leu Arg Asn Glu Leu Gln Arg Lys Asp Ala
 56 115 120 125
 58 cta gtc cag aag cac ttg aca aag ctg agg cat tgg cag cag gtg ctg 434
 59 Leu Val Gln Lys His Leu Thr Lys Leu Arg His Trp Gln Gln Val Leu
 60 130 135 140
 62 gag gag atc aac gtg cag cac aaa aag ccc gcc gac atc cct cag gcc 482

RAW SEQUENCE LISTING

DATE: 01/16/2002

PATENT APPLICATION: US/10/029,137

TIME: 11:09:32

Input Set : A:\407T-301400US.txt

Output Set: N:\CRF3\01162002\J029137.raw

```

63 Glu Asp Ile Asn Val Gln His Lys Lys Pro Ala Asp Ile Pro Gln Gly
64      145              150              155
66 ttc ttg gcc tac ctg gag cag gca tct gcc aac atc cct gca cct ctg      530
67 Ser Leu Ala Tyr Leu Glu Gln Ala Ser Ala Asn Ile Pro Ala Pro Leu
68 160              165              170              175
70 aag cca acg tgagcaaagg gcagaggcag ttggcctatg agtgggctga      579
71 Lys Pro Thr
74 tgggtgaggt tggccacaca ttcttctctg tggacttgac attttggaag aactctttgc      639
76 cagataatga gttcatttta gttttatgct cccattgaaa aattttccac tatttttata      699
78 agctgttaat ttcttgagta ctttataaca tgtctgtagc ttggataaac caagtaagta      759
80 tttttttttt gtttttagca aagttagac tgtgaatatg atgacacaga ttttttttta      819
82 tgggtgcttc gctgttttta aatttttgca tgacttttca tttttttatg tgtgtttcct      879
84 gtagtttgar ccgaaggaaa agagtatagt agcccgagaa ccaggagatg ggagttttag      939
86 tegttagcct taigataatt accccgcggt ggtgtgtaga aaagtatgta aatttgcctc      999
88 gttttaagac ttgaaactac ctcaagaaga ggaatctaat acaatatctg taatgtttcc      1059
90 agagctctca gaatgaggat ttttttgtaa ataggctcaga agacgatgga actgtcctgg      1119
92 gttagtatag taattctaca gtaggatcct taggttgatg ctgactctg tttgggggat      1179
94 gtttatattt tatgtggtgt ttactttttt tttttgacat aaaaggatat agtggggagca      1239
96 gtgatacctt aacatttcatt acattctgca gtaatgaato tg      1281
98 <210> SEQ ID NO: 2
100 <211> LENGTH: 178
102 <212> TYPE: PRT
104 <213> ORGANISM: Homo sapiens
106 <400> SEQUENCE: 2
108 Met Ala Ala Pro Leu Gly Gly Met Phe Ser Gly Gln Pro Pro Gly Pro
109 1      5      10      15
110 Pro Gln Ala Pro Pro Gly Leu Pro Gly Gln Ala Ser Leu Leu Gln Ala
111      20      25      30
112 Ala Pro Gly Ala Pro Arg Pro Ser Ser Ser Thr Leu Val Asp Glu Leu
113      35      40      45
114 Glu Ser Ser Phe Glu Ala Cys Phe Ala Ser Leu Val Ser Gln Asp Tyr
115      50      55      60
116 Val Asn Gly Thr Asp Gln Gln Gln Ile Arg Thr Gly Val Asp Gln Cys
117      65      70      75      80
118 Ile Gln Lys Phe Leu Asp Ile Ala Arg Gln Thr Glu Cys Phe Phe Leu
119      85      90      95
120 Gln Lys Arg Leu Gln Leu Ser Val Gln Lys Pro Glu Gln Val Ile Lys
121      100      105      110
122 Glu Asp Val Ser Glu Leu Arg Asn Glu Leu Gln Arg Lys Asp Ala Leu
123      115      120      125
124 Val Gln Lys His Leu Thr Lys Leu Arg His Trp Gln Gln Val Leu Glu
125      130      135      140
126 Asp Ile Asn Val Gln His Lys Lys Pro Ala Asp Ile Pro Gln Gly Ser
127      145      150      155      160
128 Leu Ala Tyr Leu Glu Gln Ala Ser Ala Asn Ile Pro Ala Pro Leu Lys
129      165      170      175
130 Pro Thr
132 <210> SEQ ID NO: 3
134 <211> LENGTH: 178

```

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/10/029,137

TIME: 11:03:32

Input Set : A:\407T-301400US.txt

Output Set: N:\CRF3\01162002\J029137.raw

```

156 <112> TYPE: PRT
157 <113> ORGANISM: Mus musculus
158 <400> SEQUENCE: 3
161 Met Ala Ala Ser Leu Gly Gly Met Phe Thr Gly Gln Pro Pro Gly Pro
162 1 5 10 15
164 Pro Pro Pro Pro Pro Gly Leu Pro Gly Gln Ala Ser Leu Leu Gln Ala
165 20 25 30
167 Ala Pro Gly Ala Pro Arg Pro Ser Asn Ser Thr Leu Val Asp Glu Leu
168 35 40 45
170 Glu Ser Ser Phe Glu Ala Cys Phe Ala Ser Leu Val Ser Gln Asp Tyr
171 50 55 60
173 Val Asn Gly Thr Asp Gln Glu Glu Ile Arg Thr Gly Val Asp Gln Cys
174 65 70 75 80
176 Ile Gln Lys Phe Leu Asp Ile Ala Arg Gln Thr Glu Cys Phe Phe Leu
177 85 90 95
179 Gln Lys Arg Leu Gln Leu Ser Val Gln Lys Pro Asp Gln Val Ile Lys
180 100 105 110
182 Glu Asp Val Ser Glu Leu Arg Ser Glu Leu Gln Arg Lys Asp Ala Leu
183 115 120 125
185 Val Gln Lys His Leu Thr Lys Leu Arg His Trp Gln Gln Val Leu Glu
186 130 135 140
188 Asp Ile Asn Val Gln His Lys Lys Pro Ala Asp Met Pro Gln Gly Ser
189 145 150 155 160
191 Leu Ala Phe Leu Glu Gln Ala Ser Ala Asn Ile Pro Ala Pro Leu Lys
192 165 170 175
194 Gln Thr
197 <210> SEQ ID NO: 4
198 <211> LENGTH: 189
199 <212> TYPE: PRT
200 <113> ORGANISM: Drosophila melanogaster
201 <400> SEQUENCE: 4
204 Met Ala Ser Asn Glu Ser Gly Gly Gly Asn Leu Met Asp Glu Phe Glu
205 1 5 10 15
207 Glu Ala Phe Gln Ser Cys Leu Leu Thr Leu Thr Lys Gln Glu Pro Asn
208 20 25 30
210 Ser Gly Thr Asn Lys Glu Glu Ile Asp Leu Glu Val Gln Lys Thr Thr
211 35 40 45
213 Asn Arg Phe Ile Asp Val Ala Arg Gln Met Glu Ala Phe Phe Leu Gln
214 50 55 60
216 Lys Arg Phe Leu Val Ser Thr Leu Lys Pro Tyr Met Leu Ile Lys Asp
217 65 70 75 80
219 Glu Asn Gln Asp Leu Ser Ile Glu Ile Gln Arg Lys Glu Ala Leu Leu
220 85 90 95
222 Gln Lys His Tyr Asn Arg Leu Glu Glu Trp Lys Ala Cys Leu Ser Asp
223 100 105 110
225 Ile Gln Gln Gly Val His Ser Arg Pro Thr Pro Pro Ile Gly Ser Gly
226 115 120 125
228 Met Leu Gln Gly Pro Gly Gly Gly Met Pro Pro Met Gly Gly Thr Pro
229 130 135 140

```

RAW SEQUENCE LISTING

DATE: 01/16/2002

PATENT APPLICATION: US/10/029,137

TIME: 11:09:32

Input Set : A:\407T-301400US.txt

Output Set: N:\CRF3\01162002\J029137.raw

```

231 Pro Arg Pro Gly Met Met Pro Gly Met Pro Pro Gly Ala Met Gln Pro
232 145 150 155 160
234 Gly Gly Pro Met Gln Pro Ser Pro His Met Leu Gln Ala Gln Gln Met
235 165 170 175
237 Gln Gln Leu Arg Met Ile Ser Arg Gln Met Pro Pro Lys
238 180 185
240 <210> SEQ ID NO: 5
241 <211> LENGTH: 1363
242 <212> TYPE: DNA
243 <213> ORGANISM: Homo sapiens
244 <214> FEATURE:
245 <216> NAME/KEY: CDS
247 <222> LOCATION: (13) (546)
249 <400> SEQUENCE: 5
250 ggcacagcaa ac atg ggc gct cca cta ggg ggt atg ttt tct ggg cag cca 51
251 Met Ala Ala Pro Leu Gly Gly Met Phe Ser Gly Gln Pro
252 1 5 10
254 ccc ggt ccc cct cag gcc ccc ccc ggc att ccc ggc caa gct tcc att 99
255 Pro Gly Pro Pro Gln Ala Pro Pro Gly Leu Pro Gly Gln Ala Ser Leu
256 15 20 25
258 att cag gca gct cca ggc gct cct aga cct tcc agc agt act ttg gtg 147
259 Leu Gln Ala Ala Pro Gly Ala Pro Arg Pro Ser Ser Ser Thr Leu Val
260 30 35 40 45
262 gac gag ttg gag tca tct ttc gag gct tgc ttt gca tct ctg gtg agt 195
263 Asp Glu Leu Glu Ser Ser Phe Glu Ala Cys Phe Ala Ser Leu Val Ser
264 50 55 60
266 cag cac tat gtc aat ggc acc gat cag gaa gaa att cga acc ggt gtt 243
267 Gln Asp Tyr Val Asn Gly Thr Asp Gln Glu Glu Ile Arg Thr Gly Val
268 65 70 75
270 gat cag tgt atc cag aag ttt ctg gat att gca aga cag aca gaa tgt 291
271 Asp Gln Cys Ile Gln Lys Phe Leu Asp Ile Ala Arg Gln Thr Glu Cys
272 80 85 90
274 ttt ttc tta caa aaa aqa ttc cag tta tct gtc cag aaa cca gag caa 339
275 Phe Phe Leu Gln Lys Arg Leu Gln Leu Ser Val Gln Lys Pro Glu Gln
276 95 100 105
278 gtt atc aaa gag gat gtg tca gaa cta agg aat gaa tta cag cgg aaa 387
279 Val Ile Lys Glu Asp Val Ser Glu Leu Arg Asn Glu Leu Gln Arg Lys
280 110 115 120 125
282 gat aca cta gtc cag aag cac ttg aca aag ctg agg cat tgg cag cag 435
283 Asp Ala Leu Val Gln Lys His Leu Thr Lys Leu Arg His Trp Gln Gln
284 130 135 140
286 gtg ctg gag gac atc aac gtg cag cac aaa aag ccc gcc gac atc cct 483
287 Val Leu Glu Asp Ile Asn Val Gln His Lys Lys Pro Ala Asp Ile Pro
288 145 150 155
290 cag ggc tcc ttg gcc tac ctg gag cag gca tct gcc aac atc cct gca 531
291 Gln Gly Ser Leu Ala Tyr Leu Glu Gln Ala Ser Ala Asn Ile Pro Ala
292 160 165 170
294 cct ctg aag cca acg tgagcaaaagg gcagagggcag ttggcctatg agtgggctga 586
295 Pro Leu Lys Pro Thr

```

RAW SEQUENCE LISTING

DATE: 01/16/2002

PATENT APPLICATION: US/10/029,137

TIME: 11:09:32

Input Set : A:\407T-301400US.txt

Output Set: N:\CRF3\01162002\J029137.raw

```

296      175
298 tgggtgaggt tggccacaca ttctttcttg tggacttgac attttggaag aactctttgc 646
300 cagataatga gtccatttta gtcttatggt cccattgaaa aattttccac tattttttata 706
302 agctgttaat ttcttgagta cttctataaca tgtctgtagc ttggataaac caagtaagta 766
304 cttttttttt gtcttttagca aagtttagac tgtgaatatg atgacacaga ttctttttta 826
306 tgggtggottt gcttggtttta aatttttgca tgaactttca tctttttatg tgtgttttct 886
308 gtagtttgat cgaagggaaa agagtatatg agcctgagaa tcaggagatg ggagttttag 946
310 tggtaggcct tatgataatt acccgcgggt ggtgtgtaga aaagtatgta aatttgcctt 1006
312 gttttaagac tttgaactac ctcaagaaga ggaatotaat acaatatttg taatgtttcc 1066
314 agagctctca gaatgaggat ttttttgtaa ataggtcaga agacgatgga actgttctgg 1126
316 gtagtatag taattcttaca gtaggatcct taggttgatg ctgaactctg tttggggtat 1186
318 gtttatattt tatgtggtgt ttactttttt ttttgacat aaaaggatat agtgggagca 1246
320 gtgatacgt aacattcatt acattctgca gtaatgaat tgaaaaaaaaa aaaaaaaaaa 1306
322 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa 1363

```

325 <210> SEQ ID NO: 6

326 <211> LENGTH: 173

327 <212> TYPE: PRT

328 <213> ORGANISM: Homo sapiens

330 <400> SEQUENCE: 6

```

332 Met Ala Ala Pro Leu Gly Gly Met Phe Ser Gly Gln Pro Pro Gly Pro
333 1      5      10      15
336 Pro Gln Ala Pro Pro Gly Leu Pro Gly Gln Ala Ser Leu Leu Gln Ala
337      20      25      30
340 Ala Pro Gly Ala Pro Arg Pro Ser Ser Thr Leu Val Asp Glu Leu
341      35      40      45
344 Gln Ser Ser Phe Glu Ala Cys Phe Ala Ser Leu Val Ser Gln Asp Tyr
345      50      55      60
348 Val Asn Gly Thr Asp Gln Glu Glu Ile Arg Thr Gly Val Asp Gln Cys
349      65      70      75      80
352 Ile Gln Lys Phe Leu Asp Ile Ala Arg Gln Thr Glu Cys Phe Phe Leu
353      85      90      95
356 Gln Lys Arg Leu Gln Leu Ser Val Gln Lys Pro Glu Gln Val Ile Lys
357      100     105     110
360 Gln Asp Val Ser Glu Leu Arg Asn Glu Leu Gln Arg Lys Asp Ala Leu
361      115     120     125
364 Val Gln Lys His Leu Thr Lys Leu Arg His Trp Gln Gln Val Leu Glu
365      130     135     140
368 Asp Ile Asn Val Gln His Lys Lys Pro Ala Asp Ile Pro Gln Gly Ser
369      145     150     155     160
372 Leu Ala Tyr Leu Glu Gln Ala Ser Ala Asn Ile Pro Ala Pro Leu Lys
373      165     170     175
376 Pro Thr

```

377 <210> SEQ ID NO: 7

378 <211> LENGTH: 13

379 <212> TYPE: DNA

C--> 383 <213> ORGANISM: Artificial

384 <220> FEATURE:

385 <223> OTHER INFORMATION: PCR primer

386 <400> SEQUENCE: 7

VERIFICATION SUMMARY

DATE: 01/16/2002

PATENT APPLICATION: US/10/029,137

TIME: 11:09:33

Input Set : A:\407T-301400US.txt

Output Set: N:\CRF3\01162002\J029137.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number
L:383 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:395 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8
L:407 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9
L:419 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10